Replies to reviewer comment #3

RC C440:
Title: Temporal and spatial scaling impacts on extreme precipitation
Authors: Eggert et al.

We thank the anonymous reviewer for the insightful comments, which we feel have helped improve the clarity of the manuscript! Our point-by-point replies (blue) to the reviewer comments (black) are given below.

Reviewer #3

1) Are these results generalizable? Some discussion on this is needed. The authors begin to talk about embedded convection and complex topography (important over narrow mountain ranges) but leave it hanging. A few sentences about whether the relationships shown here might/might not be expected elsewhere would be valuable. Will every region require it's own investigation? For example will the optimal pairs for Norway match those of Germany? Vietnam? UK?

> Since we only analyzed precipitation over Germany and are not aware of other similar studies that have been done at other climate zones, we can only speculate about this. We expect that the findings will depend on the mean advection velocity and also the orography might have an impact on the findings. We add this to the discussion part of the paper.

2) Abstract: the aim of the manuscript should appear in the first few sentences not at the end. Also it would be helpful to mention that current approaches, to say, regional modeling, do not account for spatial and temporal dependence in a rigorous way. Emphasize the results and their implications (see reviewer 1 comments on this).

> We rewrote the abstract in order to emphasis more on our results.

> The new abstract:
> Convective and stratiform precipitation events have fundamentally different physical causes. Using a radar composite over Germany, this study separates these precipitation types and compares extremes at different spatial and temporal scales, ranging from 1 km to 50 km and 5 min to 6 h, respectively. Four main objectives are addressed: First, we investigate extreme precipitation intensities for convective and stratiform precipitation events at different spatial and temporal resolutions, to identify type-dependent space and time reduction factors and to analyze regional and seasonal differences over Germany. We find strong differences between the types; with up to 30% higher reduction factors for convective extremes, exceeding all other observed seasonal and regional differences within one type. Second, we investigate how the differences in reduction factors affect the contribution of each type to extreme events as a whole, again dependent on the scale and the threshold chosen. A clear shift occurs towards more convective extremes at higher resolution or higher percentiles. For horizontal resolutions of current climate model simulations, i.e. ~10 km, the temporal resolution of the data as well as the chosen threshold have profound influence on which type of extreme will be statistically dominant. Third, we compare the ratio of area to duration

reduction factor for convective and stratiform events and find that convective events have lower effective advection velocities than stratiform events, and are therefore more strongly affected by spatial than by temporal aggregation. Finally, we discuss the entire precipitation distribution regarding data aggregation, and identify matching pairs of temporal and spatial resolutions where similar distributions are observed. The information is useful for planning observational networks or storing model data at different temporal and spatial scales.

3) P2159 L4-6: The sentence beginning "However, in many cases..." is vague and has phrases such as "...weather, respectively climate, models..." that do not make sense. The point this sentence is trying to make is important try to re-word and make it more precise.

We rewrote the part in the introduction.

Assessment of precipitation extremes, e.g. as defined by an intensity threshold, is strongly scale dependent and therefore requires specification of the analyzed spatial and temporal resolution.

Even though spatial and temporal scales are far from independent (Taylor, 1938), it is often unclear how to compare datasets directly, when their data is measured at differing resolutions. The data resolution needed by users, e.g. hydrologists or crop modelers, often differs from that at which observed or modeled data is recorded (Willems et al., 2012).

4) P2160 L10-14: It seems as though the authors wish to make a transition here from the discussion around the importance of, and challenges related to, distinguishing scales to a discussion on the physical processes governing convective and stratiform precipitation. If this is the case they should just say so, instead of the current, some- what clumsy transition paragraph.

We rephrased the sentence:

In the current study we separate the physically different processes leading to convective and stratiform type precipitation events. Using synoptic observation data, we classify precipitation events into these two types, allowing us to analyze their aggregated statistics individually across scales.

5) P2161 L4: This is actually a crucial motivation for the study and yet it is buried in the introduction. This should appear early on as a motivator and maybe even to kick off the nice literature review.

We rewrote the Introduction (see below)

6) P2161 L6-17: A whole paragraph on the pitfalls of statistical downscaling predictors but then it is not mentioned again. Is it relevant to the current study? If so, then describe why. If not, then place the discussion in the proper context or take it out.

The results of the study are relevant for statistical downscaling procedures since the change from convective extremes to more stratiform extremes, going to lower resolutions, will be a major pitfall of simple downscaling methods.

Since the study is not directly related to the choice of predictors, we shorten this part in the introduction.

7) Overall the introduction is a bit lacking. I suggest restructuring as follows: i) Start with the problem statement. Why is it important? Why should we care? ii) What have others done on this topic (literature review)? iii) What questions are still unanswered (cf. problem statement)? iv) Describe how is this study going to answer them. v) Structure of the paper

> We have completely rewritten the introduction, taking all of the reviewer points 3 to 7 into account.

New Introduction:

> The IPCC's fifth assessment report highlights an intensification of heavy precipitation events in North America and Europe (Hartmann et al., 2013), and projects further increase of extremes as global temperatures rise (Collins et al., 2013). The study of extreme events is complex due to a strong inhomogeneity of precipitation intensities in space and time. Assessment of precipitation extremes, e.g. as defined by an intensity threshold, is strongly scale dependent and therefore requires specification of the analyzed spatial and temporal resolution.
>
> Even though spatial and temporal scales are far from independent (Taylor, 1938), it is often unclear how to compare datasets directly, when their data is measured at differing resolutions. The data resolution needed by users, e.g. hydrologists or crop modelers, often differs from that at which observed or modeled data is recorded (Willems et al., 2012).
>
> The primary societal interest in extreme precipitation lies in its hydrological implications, typically requiring statistics of precipitation extremes for the area of a given catchment or drainage system, which is not identical to that of model grid boxes or the observations.
>
> Moreover, temporal scales relevant to flood risk vary enormously with area (Blöschl and Sivapalan, 1995; Westra et al., 2014): For catchments, hours to days are relevant (Mueller and Pfister, 2011), whereas urban drainage systems of ~ 10 km (Arnbjerg-Nielsen et al., 2013) are impacted at timescales from minutes to hours (De Toffol et al., 2009), and soil erosion can occur at even smaller scales (Mueller and Pfister, 2011).
>
> Areal Reduction Factors (ARF) and Intensity Duration Functions (IDF) have previously been used to describe the decrease of average precipitation intensity due to spatial and temporal aggregation (Bacchi and Ranzi, 1996; Smith et al., 1994). The capability of radar data to capture the spatial structure of storms was identified as a key factor in deriving the ARFs (Bacchi and Ranzi, 1996; Arnbjerg-Nielsen et al., 2013). A general outcome was that ARFs exhibit a decay with respect to the return period (Bacchi and Ranzi, 1996; Sivapalan and Blöschl, 1998) and a dependency on the observed region, resulting from different governing rainfall generation mechanisms (Sivapalan and Blöschl, 1998).
>
> In the current study we separate the physically different processes leading to convective and stratiform type precipitation events. Using synoptic observation data, we classify precipitation events into these two types, allowing us to analyze their aggregated statistics individually across scales.
>
> The two types physically differ in that convection is often initiated by local radiative surface heating, resulting in a buoyantly unstable atmosphere (Houze, 1997), whereas stratiform precipitation stems from large-scale frontal systems and relatively weak and

uniform up- lifting. Analyzing these two types separately regarding their intensities at different scales can e.g. be important when considering temperature changes, such as anthropogenic warming: Over large scales, the changes were found to be moderate, whereas for very small scales, it has been argued that the two processes may increase with warming (Trenberth, 1999; Trenberth et al., 2003; Trenberth, 2011; Lenderink and van Meijgaard, 2008), albeit at very differing rates (Berg et al., 2013). Using high-resolution model simulations, heavy precipitation at high temporal resolutions was suggested to increase strongly in a future climate, and a dominant contribution to extreme events to stem from convective events (Kendon et al., 2014; Muller et al., 2011; Attema et al., 2014).

In spite of their small horizontal and temporal range, convective events can cause substantial damage (Kunz, 2007; Kunz et al., 2009), e.g. through flash floods (Marchi et al., 2010).

Numerous studies have assessed the temporal and spatial characteristics of precipitation events using a storm centered, or *Lagrangian*, approach (Austin and Houze Jr., 1972; Houze Jr. and Hobbs, 1982; Moseley et al., 2013), which focuses on the storm dynamics, e.g. lifetime or the history of its spatial extent. Moseley et al. (2013) showed that, for Lagrangian event histories of 30 min, the convective type can produce significantly higher intensities than the stratiform type. As we here focus on potential hydrological applications and those addressing possible impact of extremes, e.g. floods, defining events over a *fixed* surface area and time period is more appropriate (Berndtsson and Niemczynowicz, 1988; Onof et al., 1996; Bacchi and Ranzi, 1996; Michele et al., 2001; Marani, 2003, 2005). The statistics thereby constitute averages over a defined space–time window within which both dry and wet sub-intervals may occur.

In this study, we analyze at which fixed temporal and spatial scales convective precipitation dominates precipitation extremes. To this end, we analyze two years of mid-latitude high-resolution radar data (5 min temporally and 1 km spatially), classified by precipitation types and separated into seasons (summer vs. winter) and geographic areas (north vs. south Germany). Analysis of these data over large spatial and temporal periods characterizes the statistical aggregation behavior in space and time. It can quantify the requirements on minimal model resolution sufficient for the proper description of the respective extremes. Revisiting the Taylor-hypothesis (Taylor, 1938), we contrast the two precipitation types, as to how resolutions in space and time can be compared. Using a resulting effective advection velocity, we give a simple means of quantifying effective temporal averaging in models, resulting from a given spatial resolution.

The structure of the article is as follows: In Sec. 2 we describe the data and methods used. Section 3 presents the results for extremes at different resolutions (Sect. 3.1) and suggests a method to compare the corresponding probability density functions (Sect. 3.2). We close with discussions and conclusions (Sect. 4).

8) P2163 3rd paragraph: The procedure for time steps longer than the 3hourly cloud observations is not clear.

We rephrased this sentence:

For time resolutions longer than three hours, two 3 hourly time slices have to be considered. Here we classify the precipitation event as stratiform or convective only, if the type is identified at least at one of the time slices and the other timeslice was not identified as the opposite type of event.

9) P2175 Sec3.2: A quick question on the PDF approach. Are the sample sizes for each space-time pair roughly equivalent?

As we describe in the data section, the sample size is decreasing as 1 / dx^2 dt. They are not equivalent. Where the sample size became too small, we indicate this by "missing data" in the plots.

10) Section 4: The discussions and conclusions section, like the introduction, is lacking. The first paragraph is fine but the second should make a stronger statement about how this study sets itself apart. I suggest shortening some of the text under the main headings of this section. There is too much repetition of results and not enough interpretation and contextualization. There could be a vibrant description here of the implications these results have for future modeling studies and/or observational studies. One-way to do this is to start with a bullet list of the four major findings and their main points. Then answer the questions: What are the implications of these findings? What issues or shortcomings remain? What are some potential future research directions?

The discussion and conclusions section has been completely revised, taking the reviewer suggestions into account.

New Discussion and conclusions:

Precipitation is strongly inhomogeneous in time and space. Averaging over a specific temporal or spatial interval therefore transforms the distribution function. The resulting smoothening especially affects the extreme values, as it narrows the distribution function while preserving the mean. In this study, the focus is on how such averaging affects the two synoptically identifiable precipitation types, namely stratiform and convective extreme precipitation events. Convective events are known to produce strong, short-duration and localized precipitation while stratiform events are less bursty and cover larger areas. Using synoptic observations, we separate radar-derived high-resolution precipitation intensities conditional on events of either of these two types. Unlike other studies, we here concentrate on the different aggregation behavior of the two precipitation types at different seasons and regions of Germany.

Space-time dependency of intensity distributions. We found that convective extremes were considerably stronger in the south than in the north of Germany and also showed clear seasonal differences with the highest extremes occurring in summer. Stratiform extremes showed much more moderate differences over seasons and regions.

When aggregating data temporally or spatially, we find much stronger reduction for convective than for stratiform events (about 20 to 30 % higher). These differences are larger than seasonal or regional differences that were observed within one type. This highlights the importance of distinguishing between these two types of events for example for statistical downscaling exercises. After the type separation, only the convective extremes show clear regional and seasonal differences and only in the area reduction factors. For the convective type, the strongest intensity reductions with spatial scale were found in south Germany in summer, the lowest in north Germany in winter.

Temporal and spatial scales at which shifts occur between dominantly convective and dominantly stratiform extreme events. Depending on the spatial and temporal resolution, different meteorological events will be considered extreme. We point out that this makes it difficult to compare different studies of extremes, where these extremes were de- fined at different scales. To demonstrate this, we present the contribution of convective events to the total, as a function of data aggregation, for the 99th percentile of all precipitation events.

This information is needed to identify which space-time resolutions contain comparable information about the distribution function, including the extremes. It will further help to identify at which resolution and percentile one can expect to obtain information about convective extreme precipitation events. Besides expected seasonal and regional differences with higher contribution of convective events in summer and over south Germany, we also found a clear dependency on the scale and the threshold that is used. Over north Germany, stratiform events contribute to the 99th percentile extremes only at horizontal resolutions coarser than 12 km when the duration interval is kept constant to 5 min. For a higher threshold (99.9th percentile), convective events dominate even more strongly and convective extremes consequently prevail over even larger areas and durations. **Pairs of temporal and spatial resolutions with similar aggregation effects on the extremes.** For proper choice of model output resolution, precipitation downscaling as well as bias correction, the relation between the DRF's as compared to ARF's is important. Originating from the radar data resolution of 5 min temporally and 1 km spatially, we produced sequences of aggregation, both in space and time, yielding: (i) temporally aggregated intensities for spatial scales held fixed, (ii) spatially aggregated intensity for temporal scales held fixed. Associating the respective aggregation resolution by matching identical precipitation extremes, we yield pairs of temporal and spatial resolutions, which define a curve.

The results allow, e.g., to identify pairs (Δx, Δt) of spatial and temporal resolutions for which the decrease in extreme precipitation intensities due to temporal aggregation matches that due to horizontal aggregation. In terms of the Taylor-hypothesis, the timescales can roughly be viewed as the mean duration needed to advect the precipitation pattern by the width of a grid-box (Fig. 6).

For example; if for a given horizontal grid size a higher temporal output is used, the event will likely be advected further than the size of the grid-box, leading to strong duration reduction factors. We find that for state of the art regional climate simulations, performed at a 11 km horizontal resolution, the temporal resolution needed in order to avoid stronger duration than area reduction effects, would be approximately 20 to 25 min.

In practice, in regional climate models the temporal output is often lower than the resolution computed here. It should therefore be reconsidered why many regional models do not output at sub-hourly frequency and why often only daily averages are stored.

If a model can resolve some small scale features, e.g. deep convective extremes, information can only be preserved by outputting at the appropriate temporal resolution, information lost when using lower horizontal resolutions (Fig. 8). High temporal resolution is accessible by most models already (most models have computing time steps ~ seconds – minutes) but is not routinely output at such short periods. Recording at higher frequency would mainly affect storage space, not simulation run-time (assuming efficient I/O-handling).

The pairs of corresponding grid sizes and durations defines a velocity v_{eff}, which can be used to generalize the Taylor-hypothesis to the situation where temporal scales change disproportionately compared to spatial scales (self-affinity, Deidda (2000)). For constant v_{eff} as function of spatial scale, the Taylor-hypothesis would be obeyed. However, v_{eff} of convective and stratiform extreme precipitation algebraically decreases with increasing Δx with similar exponents for both precipitation types. The main scaling difference between convective and stratiform events can be described by a constant scaling factor. This scaling factor leads to about 1.75 times higher advection velocities for stratiform than for convective events. **PDF overlap.** Changes caused by temporal aggregation depend on the spatial scale of the data and vice versa. We examine these dependencies by comparing pairs of PDFs derived for different aggregation resolutions using a method developed by Perkins et al. (2007), here defined as PDF overlap.

We find that PDF changes that were observed when decreasing the temporal resolution from 5min to 2h at 50km horizontal resolution are quantitatively comparable with PDF changes when going from 5 min to 30 min at 10 km horizontal resolution or from 5 min to 10 min at 2 km horizontal resolution.

Further we show that the PDF overlap of a certain reference resolution (we chose as an example 60 min, 10 km) compared to all other aggregated resolutions, shows a ridge with values close to 1. This ridge ranges from 5 min and 25 km to 120 min at 1 km resolution for convective type events (Figure 10c) and from 5 min and 25 km to 90 min at 1 km resolution for stratiform events (Fig. 10c). These differences can be explained by the strong area reduction factors found for the convective type. The patterns found in this analysis are very similar to, the patterns found in Figs. 3 and 4 highlighting that most of the differences found in the PDF overlap are resulting from changes in the extremes.

Technical comments:

1) P2161 L27: Change to, “Here we take the perspective of an observer capturing....”.

Sentence not included in the new Introduction

2) P2163 L15: Delete “single”

We deleted the word “single”

3) P2164 L4: Change from “is counted” to “are counted”

We changed “is counted” to “are counted”

4) P2170 L11: “Consider e.g. climate model simulation data”. There is no need for e.g. here, change to “Consider data from a climate model simulation.”

The sentence is not included anymore in the text since we reformulated parts of the chapter to make the text easier to understand.